PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/404,520

DATE: 10/06/1999
TIME: 14:29:07

OIPE

Input Set: I404520.RAW

This Raw Listing contains the General Information
Section and up to first 5 pages.

```
 1   <110> APPLICANT: Cao, Yongwei
 2         Ghodssi, Azita
 3         Hinkle, Gregory J.
 4         McIninch, James
 5         Timberlake, William E.
 6         Yu, Jaehyuk
 7   <120> TITLE OF INVENTION: Emericella nidulans Genome Sequence and Uses Thereof
 8   <130> FILE REFERENCE: 38-10(15498)A
 9   <140> CURRENT APPLICATION NUMBER: US/09/404,520
10   <141> CURRENT FILING DATE: 1999-09-23
11   <160> NUMBER OF SEQ ID NOS: 44345
12   <210> SEQ ID NO 1
13   <211> LENGTH: 5008
14   <212> TYPE: DNA
15   <213> ORGANISM: Aspergillus nidulans
16   <400> SEQUENCE: 1
17         ttcggatgac tatagagatt caaaggggtt aggttcccct gactgaatct gacactgacc     60
18         agtgcctttg cttttgttgt tgaaacacca tctatacagc cattgctgat ctcccgtcca    120
19         aaccacacca acaacgccaa cctactttgg ctcaacacca cgcgctggtt tctgctcaga    180
20         attacttgga aataagccta cttagattta agtaccgtcc ccataccctg gtcaatcctg    240
21         agcaggatcg gccgactggt tgcaacagaa cacttcgccc aagagccaac ggaaccattc    300
22         gctttggccg aagattgtgc catacaaagc aattagacct taactccacg attctattgt    360
23         tcccggggcc cagtgggata gggttacctt cacagtaaac cgcccatgga caataaccca    420
24         aaataatcca gttatgcata taatagcttt aacaggcagt gatctgcatg gccaaataaa    480
25         atactgtatt ctaatatagt attataaact atttatgtat accaaaataa tcaaaataca    540
26         gtaatatacc tactcagata gcttgccaac cagcgagttg ctccgccggg ctttggggca    600
27         gccaaaacat ccaaatccaa tggattatta catggtctaa cccaacccaa tccttggcgg    660
28         gtcggggggg gttggggcgg gtttaacaag tctagtctag attgaagaag gcacagtttg    720
29         caagctgctg ggtgggggggt ccatggcatt agtatgctgt gtggacccgc agaggcggga    780
30         gcggttgcgg acctggagtt gtgccacatc cagcattagg ggtgcagtgc cgtccgggtt    840
31         aggtttcaac ccactctgca atgaaaacaa caccactcaa tttgtgtgct accgcgcggg    900
32         tttaaaaaag acattagaat acgtacattt cataagtttt ttgatataat gcattcttac    960
33         gatattttat gcaggttagt atgtttctac gcatttttca cccgcacggg tctcctatca   1020
34         ggaaccgtag cccgcgcgga ctgccaatta tgcgactctc cccgcaccgc tgcgggttgc   1080
35         ggtgcgcgtt agcaacccta tcgagcagca ggtgcattgt atagtcttgg ccttcctaga   1140
36         taccagacac ttcgtaatta atacgagtta ttgggcatgt gtcacgctag ctgcatcaca   1200
37         tgacttgctt ttagggcgcc cgatcatttg gactgccacc acctgtaagc tggcctccta   1260
38         acattaatga gtaggcaaac tcatcctctc taacagataa tcataaggcc atattgaatg   1320
39         agcaaaggtg agtaacccat ttgatacact gatgccattt ggacttcagc agttgtctaa   1380
40         tctttcgcag ccccaaatat ataaacacat acaccactct acgatgatgc tccccgccaa   1440
41         atcgactctc cttgccctcg cctccttcat atctctcgcc gtggctgtcg atgtcggtgc   1500
42         accgtgcgac agctcaagag agggtgtata ctggtgtgca aacgaccaag ctaccattgc   1560
43         tgtctgccat cttggagcct ggcatgagtc ttctaaatgc gggattggtt gctgcgcatg   1620
44         gcattgggtt aataatttgc catattgctt ttgctaggca aagcgttgat ggcccctagt   1680
```

```
45    gctagagatg atggagttag ggcggggagg ccagtaatta tctggactgg gatgaggggt 1740
46    tgcagtgggg tttattgaaa cctgcagata aattcacggt gatatgagga agctggcaca 1800
47    gaataggatg aactaagtaa ttagtgtctg agggagcgag tttatggact tagttagaga 1860
48    acagaaaggt ggtagatata cacgtgtcat cttacatagc cccttgctag gggatgtaat 1920
49    cagatggatc atcctatctg tcatagaata tgctcaaagg ggaaatggta gaaggaaact 1980
50    gtatctataa caggcatggg cattattgtt ttcaactggc tgtgctgtat tcacaagtat 2040
51    cataagcttg aagagaggta tgggattgcc ttcctctccc tctatataca tgctcctaag 2100
52    cttccatcgg gttttattca acttcctaga atgttctaac ttgtctgcca ttcaagtact 2160
53    atggtagggc ttcctgtagc ccagcccagg gacaaccgtg catggtacgg caacgcgctc 2220
54    tcaaagtcaa ggactaacta aattgactgc acaaacacct ctctctgcac cacgattttc 2280
55    tgcgccctcc ctaaaaaccc caatttcttg cgcatgccgc tggtctgagt cccaaggccg 2340
56    cagctcctcg cgtgaagagt aggacccctc gtaccctggc tctggcagct ccgacctacc 2400
57    cccctccaa actcaacact tcactacgac cctcatctgt cggatagaca ccccgcaccc 2460
58    tgacccagac acacttcagc agcggcacat acaacaccgg gcacgcggtg acctggtata 2520
59    tgcccgcctc gactgtcgcc tacatgagga tgatagcccc ggaccaggtg ttattcatga 2580
60    cggcactggt ggacgatatt ttcgtcaggg agataattgc cgcgttaaga ccactgtgag 2640
61    agtgtcacta aagcaaaaca aaggcaggga gacgaaatgg tggatggaca tacatattcc 2700
62    ctacgccaaa cgccgccaaa ataccaatct tcattcgcag cctcatctgc agcctcgcca 2760
63    cgtgcgggat cggaagcgcc agcctcctga cgtcgcaaac gatattcacc acccgcgttg 2820
64    tgcgtaacat cgtgtttatg gagacgcact ggcccggcac gcccgcgtcc cagagtttgg 2880
65    tcagcggcct gcacacgaca cttaggagga ctagcgctat cagcacattt gcgagaatga 2940
66    ttgcgccgag cgcgtagcag gttatttttg ttgggcggtg gatgccgaag gtttggaggt 3000
67    ggaggcttag gattgcaagg cgcgggggca cgacggcgag tgcgtaaagg tacgagattg 3060
68    cgactaggaa gattgcccca ggtttagctg gtaagaggct ctgcgacatt agtgcaggcg 3120
69    gacgccggcg tatttggtgt cagctgttcg gggccgtccg tggttagcat tgaggtacct 3180
70    gcctatctaa tggtcggggt ctgggtatct ggctgttgca tcctgggagg tccaaaggtt 3240
71    gctcgaatag tctctacagc tcttcaagct tattctgtct ttgcaaggtc agatttgctg 3300
72    gcttcgcctc cttaggctta ggtatagaag ctgtagccaa ctcaactgct aattctgttc 3360
73    gcagttgctt gctaggattg tctctcttga tacgactatc ttttcggatt ctcttgtctg 3420
74    tctcgtccgt cttgtatatg gttgccccgg tcgtctcagt gggctctatc tcgaccgcct 3480
75    cgctcagccc tggtgccgtc tcggtctcat actcctctct attggttcat caaatgcccg 3540
76    agtacgttcc tactcggccc cgctaggtct ttcgtattct ctcggctcga tgtaacatct 3600
77    cccggttcgt ggatttcgtc ttcgagccag cggcatccgc cttgttcggt tctaggttcg 3660
78    ttgttctata gctcgatcga ggatagatag gtaagttgta caggcctggc caccgtcaca 3720
79    ggagcaaggt tcaacgggcc acgtacagtt ccgatcttct gctgaattgg gtgagtacat 3780
80    tatgtctcac acattccccc tagatgctag gcctgctggt taggctagat gtagcagcct 3840
81    gggctcattg ccaatgttgt ggattagact tgttaaacca cgggttgggg cgggttttca 3900
82    ggcctagctg atccgcccac gcgggttttg gggtgggtta ccttcacagt aaaccgccca 3960
83    tgggtttagc aaataattct aacccaacct aaataaccca aaataaccca gtaatgcata 4020
84    tcattactct aatagataat aatctacata gttaataaaa tactgtattt aaatactgta 4080
85    ttataactat ctaagtaaga aaatataatc taaatacagt aatataccta tttagatatc 4140
86    ttggcaaccc agcaggttgc tccgccgggc tttggggcag ccaaaaatat ctaaaaccta 4200
87    atagataatt agaaggtcta acccaaccca tttcttggcg ggtcggggcg ggttggggcg 4260
88    ggtttcgtgg gttgggttta acaagtctag taatagactg ctaacttctt agaatgttac 4320
89    cggcaccttt agtccatata ttaggtgtat taatgataca actaataatc tgactaaata 4380
90    ttactgtagt tctattcaag gctatgcagt gcttgatctt cagttcttag aggatctttt 4440
91    taataaagat attctaccag cagctgagga atgtggtgtc tttgaactat tatctgtagc 4500
92    gggccctgcc cggaagtagt gctcgaggcc gtcgacggcc actcgcttca cgtgaccttt 4560
93    ccgtcacatg acccatcggt aagctggcgc atatctagat gtttctatcc tgttgatctt 4620
94    gtttcttctc ttttcttctt tctttagcgt gttcttctgt gatgatccta cctgtaggat 4680
```

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/404,520

DATE: 10/06/1999
TIME: 14:29:07

Input Set: I404520.RAW

```
 95            aattaatggc attagcccgt aacattagaa cacgctcaca accactacta caacaagacg 4740
 96            tcggacagcg gacattggac aacagacatt ggacaacgga cattggacaa cggacactgg 4800
 97            acaacggaca ttggacaacg gacattggac aatggacatt ggataacgga cattgggcag 4860
 98            tggacattgg aaacagacat tagtaacgga cattggtaac ggacatcgga agacggagac 4920
 99            tggagagtga cttttggaca acggtttgaa caaccagtcg gacaatacgg acagcgaacg 4980
100            tgacaccgga cgcgacactt ataccgaa                                    5008
101    <210> SEQ ID NO 2
102    <211> LENGTH: 1209
103    <212> TYPE: DNA
104    <213> ORGANISM: Aspergillus nidulans
105    <400> SEQUENCE: 2
106            ggctgatatc tggcctagtt tttgtagcta gatactgcaa gctgccagta ccgtgctggt   60
107            attctttgat attgattttt ggatctggag catcatcgat aactgttcca ggatcaaggg  120
108            gtgatttggt aggatgggca ttatgaagcc taaaagagtt gataatatta ttaatttgat  180
109            cctcctgtga caggtggatg cctgtaggcg ttgttttgat atttatgcct agatatcgct  240
110            tgaattccgg cgcttcttta atttcaaatc gggatttcaa gctttcgagc gcttgtacgg  300
111            cgttttgggc cttttcacca acaatcatga agtcatcaac atgggtggtg atatagagct  360
112            tttctgtaac atgtacgaat aggcctgggt catgcgggga aaccttaaat cctatagatt  420
113            caagatatgc tcgtagatca ttgtaccata ggcgcgcgga cggggttagg ccatatagtg  480
114            attggcggag caggcaaact agggtccctg gctccccttc agcgaatcca agtggttgat  540
115            gcatataaat cggttggcct tttaactttg tatttagaaa tgctagaact gtatctgcct  600
116            gaataatatg ccagcctttt tgggcgctga caccaaacaa gatttgtgaa gtaataggat  660
117            caacaactgg ggcgtatgtt gctccttcaa attctgattt atcaagaaga ttccccctga  720
118            ttacctattg tgctttgtaa cggatagagt tgtcaggatt aagcttcttt ttataaaccc  780
119            atcttcctgg tatggcgcgt gcgccagcag gcatgtcaga ttttcagatc aaatcccaac  840
120            attgcttctg tttcaaggta tttatctctt cttgcattgc tgcatgccat tgaactgatt  900
121            cagggccttt aatagcctca ttgtacgatt tgggcttagg tagaggtgat tcaacagatt  960
122            ttaggatagc aaatcctttg gcaggtcggt atagctcttc tccggaggta atatctgata 1020
123            tatcctgtat aagaaggtct tcattggttt tgggttctgc agcttcggca gtttcgccag 1080
124            gtttgacagt ttcgagagct ttaccagctt cgggagcttt ggcagcttcg tacggcttgc 1140
125            tcagtggcgg gttaaactcc tccttgatag agatatcact ggtatcacgt agttggttga 1200
126            tttcagggt                                                         1209
127    <210> SEQ ID NO 3
128    <211> LENGTH: 1567
129    <212> TYPE: DNA
130    <213> ORGANISM: Aspergillus nidulans
131    <400> SEQUENCE: 3
132            gaaccaggtt gagggcatgg cgacaaccaa gatcaggcca tagcctgtga caaccactat   60
133            caaatacatc aatactaagg ggtggccaat accctccacc attatcttta aaggaaaagt  120
134            ctatatagag ggatagttta ttgaatgcat ggtccctggc agttagagga ttgaaataag  180
135            caccaatagt tagactatag atataattgg ctttttgctgg cttcaaaata tcttcattcc  240
136            agcgacaaat aagcgtacaa ggggggggata ttggctgctt atctcacctc ctaaacccctt  300
137            tggatgttgg ttgttgcaga cctttaaaga aggcatatag caagcttatt gaggagaaat  360
138            gacgcctggg atataactat attaacatgc tagatttctt gaaattttac ctagcaggtt  420
139            accaggaagt ctttactaca gagaatattt aaagcagatt caaagtaact aggatactac  480
140            cttctgatct catggtagtg cttgaaaagt caaatatcag cctgggtact ctaacccccc  540
141            ctccaagcca agggggtgct tcaatccctt cttctcagct tggtacgcct tatactgtgt  600
142            actatgtata ttgaaaagtc tcttcagtta aaaggatgct tgagagaagg ttaaaagtcc  660
143            cctaactccc tccaaaagaa gtcttagata agttggtaaa agggtgtaag ttggctatct  720
144            ataatacaac ttactagcaa gggaaaactg tggtctctac tcagccatag agaataacag  780
```

PAGE: 4

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/404,520

DATE: 10/06/1999
TIME: 14:29:07

OIPE

Input Set: I404520.RAW

```
145        acaggaaaag agttattctt aatactagat aactcctata caatgtcttc cattccagga   840
146        agccagggac ctaatttaga caagaaataa ctaaatacag gcaagatagg ggggtgctgg   900
147        cggaggtgcg cctcaaactt tagatgcacc taaatatact ctaccaacat gttttagaat   960
148        ataatattca gggctataag aggaccagct gtcctactca ttatagaatt tagtttattt  1020
149        taatttgaat cacttttagt tgacgcacag tacttcagag tcgagcaaga ataggttttg  1080
150        aagggaatct cacggatcac ccggatttta cgttaaagcc tagaggataa tgcaactgaa  1140
151        agcatactca ccaaaataat tgttaacctc tgtttaagaa ttaatactgc cagcgtatat  1200
152        cagacctatc cttatttagt acccaagtga agaattgaag tataattcga gtattatagg  1260
153        taaaagaatg aatagttagg ttaataaacc gatttctata tttaggcagg ataggcagat  1320
154        tggaatagtt aaaactagta ttttttcact atttgggtgt ttatagccta gaatcattta  1380
155        tccctagttc tctcgttttg tgttttattt atctgtccat tactaaatca atatcatttt  1440
156        tattgttttt cctctttatt attatttggc gttatgttct ttaaattact tttagttcat  1500
157        ttattttcca ttttttttttt attttcttct cttttcatag ttactcttct tcttccatcc  1560
158        atcttct                                                           1567
159   <210> SEQ ID NO 4
160   <211> LENGTH: 408
161   <212> TYPE: DNA
162   <213> ORGANISM: Aspergillus nidulans
163   <400> SEQUENCE: 4
164        gttgatggag gtactctctc agctgtggac aaccagcggc ccaacatcac cctggaaggc    60
165        aaccaacgcc acctttattt attctctact ccttcctact ctaaacacat ttatccggag   120
166        tctgctgtct gccattgtgg ggctaccgga ggatgacatt accacgtcca tggccagcga   180
167        tatcctgacc tcaacttctc cgatggcaac cctagtcctc tcctttatct cctcttagag   240
168        gctcgcacat atctcatctt gactcccgtc actcatggcc ctcgctctct tctgcgcgcc   300
169        attcgtcaga tcccaactgc caactacatc gcaagtactc cgttgattcc aaattgtatc   360
170        ttctctcgac ccgtcctaca ccgtcctcct ggaatttatt tagcttca               408
171   <210> SEQ ID NO 5
172   <211> LENGTH: 4602
173   <212> TYPE: DNA
174   <213> ORGANISM: Aspergillus nidulans
175   <400> SEQUENCE: 5
176        aggagtcagt attgcaaggc tacaaagata atgaacaaat actatattga tattgcaggg    60
177        gggtgaagca taattagacg gtaaggatga agaaagcccc tccgggtatc aggaatattc   120
178        caatgtttct tgctgattac agacaatgca acctgatatc aatgggttta tacccttcaa   180
179        accagagata gggctattcc cacctttcag ggatggctga attatttcaa aaaccaagga   240
180        tatagcccac cagctttggt acgaagtgat cgcgaattta taactgaaaa tgtcaagaaa   300
181        ctatgccaaa actatagcct attttgggaa ccaaccgctg tagactcccc atggcaagat   360
182        ggcgtcagtg agcgcggaat acagacagtt ttgcagtata cacgggcaat gttatataac   420
183        tcaggattac catgctggct atggccgctg gctttacaga cagctgttta tcacataaac   480
184        tgattaccta taagggtcct cttgtacaat gatcgacggc ctatggcact aaccagcgat   540
185        ccagaaattc agccatgtgc ccattttacg ccctattctg cctggatcaa tggcgatgcc   600
186        aatatcaagc atcttgtcaa atttgggtca cctgcttgga tgcacctaca tggagcttca   660
187        aaatatgcta gcaagccaac cagcaagatc gacccaaagg ctaaaaaggt acatgttgtc   720
188        ggatatcaag gctcccatat ctatgttatt tgggaccccg aaatcaatca actccgtgac   780
189        accagtaata tcttgatcaa ggaagaattc agccctccac agaacaagct gtacaaagcc   840
190        accaaagctg ccaaaactat taaagctacc aaagctgcag aaccaaaaac caatggaaaa   900
191        ccttcctata caggatatat cagacattac ctccagagaa gagctatacc gacctgccaa   960
192        aggatttgct atcctaaaat ctactaaatc acctctacct gagcccaaat tgtataatga  1020
193        ggctattaaa ggccctgaat cagcctaatg gcatgcagca atgcaggatg agatcaatac  1080
194        cttgaaacaa aaataatatt aggatttggt ctgaaaatct gacatgcctg ctggcacaca  1140
```

```
195    tactatacca  ggaagatggg  tttataaaaa  gaagcttaat  cctgacaact  ccatctgtta  1200
196    cagagcatga  taggtaatca  gggggaaatc  ttcttgataa  atcagaattt  gaaggagcaa  1260
197    cgtacgcccc  agttgttgat  cctattactt  cacgtatctt  ttggtgtcag  cgcccaaaaa  1320
198    ggctggcata  atatcaaacc  agatgcagtc  ctagcattcc  taaatgcaaa  gctaaaaggc  1380
199    caactgattt  acatgcacta  accactcaga  ttcgcgaagg  ggagccggga  accctagttt  1440
200    gcctgctcca  ccaatcactc  tatggcctaa  ccctgtccgc  gcgcctatgg  tacgatgatc  1500
201    tagcatatct  tggatctata  ggatttaaaa  tctccccgca  tgacccaggc  ctattcacac  1560
202    atactacgga  aaagctctct  atcaccaccc  atgttgatga  cttcatgatt  gttggtgaaa  1620
203    agccccagaa  cgccgtacaa  gcgcttgaaa  gcttgaaatc  ccgatctgaa  atcaaagaag  1680
204    tgcctgcatt  caagcaatat  ctaggcatga  atattaaagc  aacatctact  ggcatccacc  1740
205    tatcacagga  ggatcagatt  gatgacctcg  tcaactcttt  caggctccat  aatgcccatc  1800
206    ctaccaaatc  accccttgat  cctggaacaa  ttatcgatga  tgctccggat  ccaaaaatca  1860
207    atatcaaaga  ataccagcgt  ggtactggcg  tcttgcaata  tctagctaca  aaactaggcc  1920
208    agatatcagc  tgagctgcct  ccgtcgcttc  tctctgcatg  tcaaccagaa  gattttccaa  1980
209    atctggatac  aatagctctt  gcgggtgagc  cggtgccaca  ggggttagcc  gacgcttggg  2040
210    tgagtagaca  cttactcaat  ctatatgggc  ctgcagaggt  acgcacctct  gtttcttttc  2100
211    cagaggaaaa  tcattaacgg  gacatgacag  tgcggcatga  tctcaaccgg  tacacgcctc  2160
212    ttccctcgtc  aaaaaatcac  tatcggcctg  cctctgccgg  gaatgaacgc  atatttattc  2220
213    aaccaccgcc  gatgcccagt  tcccgctggg  gttatagggg  agctgtatct  ctctggggag  2280
214    caagtgacgc  cagggtactg  gaatttcaag  cataatgcgg  ctttcacaca  cgaccccttc  2340
215    tctccaggcc  agattatgta  tcgaaccggc  gacctggcgt  actgggactc  cttcaagaac  2400
216    cttgtatacg  tccgcagagt  ggacaaccag  atcaaagtac  gcggttaccg  tgttgaactt  2460
217    gaggaaattg  agcacgcgct  gggcagagcc  gatagccatg  ttcggagcgc  agcggctatc  2520
218    gtcgtagata  atgtgcggat  aatcgcattc  gtcgcgccag  aaacagtcga  cgcctctgcc  2580
219    ctccgtcgac  aactggttac  cttgctgccc  aaatatacca  ggcccaccga  gatcttcgca  2640
220    ctcccatctt  taccgacgtc  tgcgaatttc  aagatcgacc  ggcgggtgct  gctagacctg  2700
221    gcgagaacaa  acaaaggcca  aggtgaactg  ccatctacac  cgacagagac  tcttattgcg  2760
222    ggaatctgga  gaaatttgtt  acagtctcac  ttagacccga  ctcgggaaat  ccaccgcgat  2820
223    gatgatttcc  tgggcattgg  cggaacctcg  ctgctctcaa  tcgaggcggc  gagaagcata  2880
224    tctaccgcac  taggccatca  gattcctatc  tcggtactca  ttcgggagac  catttttgtct  2940
225    gacctggcag  gcagggtcga  tgagtatagc  aaaaccgagc  acggcaccac  cggctcgttt  3000
226    gcatcctatc  tcaaggcaaa  ccctactcta  gccgaggacg  taactacacc  ttcctacctc  3060
227    gaagaaggac  tcttccatct  ctactcacaa  acgagcacga  aaagtgcatt  tcgcgtcgct  3120
228    gttcagttcg  tcgctaacgg  gattgttaac  catgaggctc  tatgcgatgc  ctttatagcc  3180
229    atcatccgcg  aagacccat   cctccgagca  cgctattccg  ccgaagaagg  tgttcttcgc  3240
230    cggacgatta  gtgaacaggt  ttcccctccc  aaattttacg  ttgggagtgc  gattaattct  3300
231    gagcagcttg  aggatcttgc  taacaagccc  ttcaaccttg  gatcggatca  gcttctgcgc  3360
232    gttgtgatct  ggaagagaga  tacaaaaacg  acagttatta  taattgttgt  ccaccatatt  3420
233    atcacggata  aagcttcaat  tgctcttatg  ctacaggctt  tgacccataa  ctaccatgct  3480
234    gcccttggaa  gactttcagc  aactcataac  ccgtgcacca  gactatccga  gtaccaaggc  3540
235    aactatataa  cgtgggcgca  gtggctgtgc  gataccggaa  gcaccagagg  cagcacagaa  3600
236    gcaacgcgca  aaagggaagt  gttctggaaa  tcctatcttc  gtgatataac  cctattgccg  3660
237    gctctttctt  cactctctag  caagactgcc  tgccacgagg  gcactcaccg  ctccaatctc  3720
238    atcccatacc  acgcaggaaa  gccgtactca  cagctcgcca  tcgcaacaac  agcgctttct  3780
239    ctccatagcg  cattccaagc  gactgatatg  gtcttggctg  tcccatttgt  caaccgcgac  3840
240    gacgaagcca  cggccacaac  cctcggactg  ttcgttgatc  ggttacctat  ccgcctaaac  3900
241    ctgaataagt  acttggagtc  tgatagcata  ctccaccaga  ataaagacac  attcgtcgca  3960
242    gatgtatcag  atcagataac  tggtgctgtg  gataattacc  tgccttatag  gcacattgca  4020
243    aaactcctgt  ctgacacgac  cacggatacg  gacagccttc  ctgctgaacc  cttcttcgcc  4080
       gcgatggtcg  tctaccattg  gcctctgatg  cgctagagaa  cagttttaac  ctctaactcg  4140
```